

# Empirical correlation of PGA, spectral accelerations and spectrum intensities from active shallow crustal earthquakes

Brendon A. Bradley<sup>1,2,\*</sup>

<sup>1</sup>*Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand*

<sup>2</sup>*Institute of Science and Engineering, Chuo University, Tokyo, Japan*

## SUMMARY

Empirical correlation equations between peak ground acceleration, spectral acceleration, spectrum intensity, and acceleration spectrum intensity are developed. The correlation equations are developed for shallow crustal earthquakes using the Next Generation Attenuation (NGA) ground motion database, and four of the NGA ground motion prediction equations (GMPEs). A particularly novel aspect of the present study is the explicit consideration of epistemic uncertainty in the correlation equations due to both the adopted ground motion database and GMPEs. The resulting correlation equations enable the joint consideration of these four ground motion intensity measures in ground motion selection using frameworks such as the generalized conditional intensity measure approach. Copyright © 2011 John Wiley & Sons, Ltd.

Received 19 September 2010; Revised 15 January 2011; Accepted 18 January 2011

**KEY WORDS:** correlation; spectrum intensity; acceleration spectrum intensity; active shallow crustal; generalized conditional intensity measure (GCIM)

## INTRODUCTION

Earthquake-induced ground motions are complex excitations resulting from a complex rupture process and seismic wave propagation. In order to simplify such complex ground motions, they are often quantified via a scalar ground motion intensity measure (IM). The variation of such IMs with variables such as source magnitude, and source-to-site distance, formally quantified by ground motion prediction equations (GMPEs), provides both insight into ground motion phenomena and are also utilized frequently in engineering design and analysis.

In emerging performance-based seismic design procedures, seismic response analysis employing ground motion time histories is becoming more prevalent. In such cases one is often faced with the challenge of determining which such ground motion records (either recorded or generated) should be used in assessing seismic performance. The challenge of ground motion selection is primarily a result of the fact that GMPEs provide only predictions for a single IM. By definition, IMs represent a highly simplified representation of the ground motion. Hence, it can be generally appreciated that the use of a single IM will provide an insufficient representation of the severity of a specific ground motion for a specific structure [1–4].

The generalized conditional intensity measure (GCIM) approach [5] provides a probabilistically robust framework in which any number of ground motion IMs can be considered in ground motion selection. A key component of the GCIM framework is the availability of equations which provide

\*Correspondence to: Brendon A. Bradley, Department of Civil and Natural Resources Engineering, University of Canterbury, Christchurch, New Zealand.

†E-mail: brendon.bradley@canterbury.ac.nz

the correlation between (the logarithm of) various ground motion IMs for a given earthquake rupture. Empirical correlation equations are already available for some IM combinations. For example, equations between (pseudo) spectral acceleration, SA, values at different vibration periods have been developed by Inoue and Cornell [6], Baker and Cornell [7], Baker and Jayaram [8], and equations between SA, peak ground acceleration (PGA) and Arias intensity ( $I_a$ ) [9] were developed by Baker [10].

This paper provides empirical correlation equations between geometric mean PGA, SA, spectrum intensity (SI) [11], and acceleration spectrum intensity (ASI) [12]. The correlation equations are developed using the Next Generation Attenuation (NGA) ground motion database [13] and are applicable for ground motions from active shallow crustal earthquakes. Explicit consideration of, and predictive models for, the epistemic uncertainty in such correlations are developed due to the adopted ground motion database and GMPEs. The correlation equation developed between PGA and SA, in particular, is compared with other previously presented equations.

## INTENSITY MEASURES, PREDICTION EQUATIONS, AND GROUND MOTION DATABASE CONSIDERED

### *Intensity measures examined*

The severity of a specific ground motion for a particular structure is in general a function of the amplitude/intensity, frequency content, and duration of the ground motion. This paper is limited to the development of correlation equations between several intensity measures which represent the amplitude of a ground motion at various frequencies. As such, the specific intensity measures considered are: (i) PGA; (ii) (pseudo) spectral acceleration, SA, for periods from 0.01 to 10 s; (iii) spectrum intensity, SI; and (iv) ASI.

PGA has been used extensively in the past because of the simplicity with which it could be measured. SA (which is uniquely related to spectral displacement,  $S_d$ ) is the most commonly used intensity measure in both seismic design and the selection of ground motion records for seismic response and performance assessment [14–17]. SI and ASI, which represent the average intensity of a ground motion for moderate and short periods, respectively, have also been adopted in various forms of seismic response assessment, e.g. [1, 2, 18]. The averaging (i.e. integral) nature of SI and ASI offers the benefits of good prediction for general moderate, and short-period structures, respectively [2], as well as a reduced prediction standard deviation in comparison with spectral accelerations at single vibration periods [19, 20]. The number of ground motion IMs considered herein is limited because of scope and is not intended to imply that other IMs are not important in seismic hazard analysis, seismic response analysis, or ground motion selection.

### *Ground motion prediction equations adopted*

The predicted distributions of the aforementioned intensity measures for each ground motion record were computed using four of the NGA GMPEs: Boore and Atkinson [21], Chiou and Youngs [22], Campbell and Bozorgnia [23], and Abrahamson and Silva [24]. These four GMPEs are herein referred to as BA08, CY08, CB08 and AS08, respectively. The four NGA GMPEs provide predictions for both PGA and SA. These GMPEs can also be used to predict SI and ASI using the analytical equations developed by Bradley *et al.* [20] and Bradley [19], respectively. Herein, for example, a prediction of SI using the Bradley *et al.* [20] analytical equation and the BA08 SA model is simply referred to as a BA08 SI prediction.

### *Ground motion database*

Ground motion records from the NGA database [13] (<http://peer.berkeley.edu/nga>) were used to obtain empirical values of PGA, SA, SI, and ASI for use in the subsequent correlation analyses. A specific subset of the NGA database was obtained by starting with the data set of 3297 records used by Chiou and Youngs [25], and then removing those records associated with aftershocks, to

obtain a data set of 1842 motions. While this subset is not the same as the individual subsets used by the various GMPEs developed from the NGA database, all of these such GMPEs are applicable for a range of predictor variables (i.e. magnitude, distance, etc.) which encompass all of those records used in the present study. This ensures that the point estimated correlation coefficients computed subsequently are unbiased.

Empirical ground motion IM values and the predicted IM distribution (using GMPEs) for each record were only computed if the IM was within the usable period range of the record. For example, if the maximum usable period was 2.0 s then no spectral ordinates above 2.0 s or SI (which is based on SA from 0.1–2.5 s) would be computed.

## METHODOLOGY FOR CORRELATION COEFFICIENT COMPUTATION

### *Point-estimated correlations from ground motion data*

Prediction equations for ground motion IMs have the general form:

$$\ln \text{IM}_i = f_i(\text{Rup}_k, \text{Site}) + \varepsilon_i \sigma_i(\text{Rup}_k, \text{Site}) \quad (1)$$

where  $\ln$  is the natural logarithm;  $f_i(\text{Rup}_k, \text{Site}) = \mu_{\ln \text{IM}_i | \text{Rup}_k}$  and  $\sigma_i(\text{Rup}_k, \text{Site}) = \sigma_{\ln \text{IM}_i | \text{Rup}_k}$  are the predicted mean and standard deviation of  $\ln \text{IM}_i$ , which are a function of properties of the earthquake rupture and site of interest; and  $\varepsilon_i$  is the (normalized total) residual (assumed to have a standard normal distribution).

By rearranging Equation (1), the residual,  $\varepsilon_{i,n}$ , can be interpreted as the number of standard deviations a particular ground motion,  $n$ , has  $\ln \text{IM}_i$  above the predicted mean:

$$\varepsilon_{i,n} = \frac{\ln \text{IM}_{i,n} - \mu_{\ln \text{IM}_i | \text{Rup}_k}}{\sigma_{\ln \text{IM}_i | \text{Rup}_k}} \quad (2)$$

As previously mentioned since a single IM represents a significant simplification of a ground motion then the use of multiple IMs (i.e. a vector of IMs) may be desired. The probabilistic characterization of a ground motion using multiple IMs must account for the fact that the individual IMs are correlated. For example, if a ground motion has a higher than expected 0.2-second spectral acceleration (i.e. SA(0.2)), then it is also likely that SA(0.3) will be higher than expected because of the similar vibration frequencies which these intensity measures are sensitive to. Furthermore, it is also likely that such a ground motion will have a higher than expected ASI, because of the fact that ASI is the integral of SA from 0.1–0.5 s.

From the linear relationship between  $\ln \text{IM}_{i,n}$  and  $\varepsilon_{i,n}$  in Equation (2) it follows that the correlation between the logarithm of two IMs, for a given earthquake rupture, is equal to the correlation between the normalized residuals, i.e.  $\rho_{\ln \text{IM}_i | \text{Rup}_k, \ln \text{IM}_j | \text{Rup}_k} = \rho_{\varepsilon_i, \varepsilon_j}$ . Herein, for brevity, this correlation between two intensity measures,  $\rho_{\ln \text{IM}_i | \text{Rup}_k, \ln \text{IM}_j | \text{Rup}_k}$ , will simply be referred to as  $\rho_{\ln \text{IM}_i, \ln \text{IM}_j}$  (where the conditioning on  $\text{Rup}_k$  and the same location is implied). Hence, using the values of the IMs computed for each ground motion record, the correlation coefficient between two different IMs (at the same location due to a given rupture  $k$ ) can be estimated using the (maximum likelihood) Pearson-product-moment correlation coefficient [26]:

$$\rho_{\ln \text{IM}_i, \ln \text{IM}_j} = \rho_{\varepsilon_i, \varepsilon_j} = \frac{\sum_n [(\varepsilon_{i,n} - \bar{\varepsilon}_i)(\varepsilon_{j,n} - \bar{\varepsilon}_j)]}{\sum_n [(\varepsilon_{i,n} - \bar{\varepsilon}_i)^2] \sum_n [(\varepsilon_{j,n} - \bar{\varepsilon}_j)^2]} \quad (3)$$

where  $\varepsilon_{i,n}$  and  $\varepsilon_{j,n}$  are the normalized residuals of  $\ln \text{IM}_i$  and  $\ln \text{IM}_j$  from ground motion record  $n$ ,  $\bar{\varepsilon}_i$  and  $\bar{\varepsilon}_j$  are the sample means of  $\varepsilon_{i,n}$  and  $\varepsilon_{j,n}$  and  $\sum_n [ ]$  represents summation over the number of ground motion records (i.e.  $n = 1 - N_{\text{record}}$ ). It is noted that the determination of the correlation coefficient using the total residuals directly in Equation (3) is not strictly correct, given that ground motions from the same earthquake are recognized to be correlated [27]. However, the determination of the correlation of the total residuals based on the correlations between the inter- and intra-event residuals is a non-trivial task, and beyond the scope of this work. Therefore, the

total residuals were used in Equation (3) directly based on the knowledge that the bias incurred is likely small [8].

#### *Uncertainty in the correlation coefficient*

As can be seen from Equation (2), the residual for a particular  $n$ th record,  $\varepsilon_{i,n}$ , is a function of the GMPE used to estimate the mean,  $\mu_{\ln IM_i | \text{Rup}_k}$ , and standard deviation,  $\sigma_{\ln IM_i | \text{Rup}_k}$ . Therefore, the correlation coefficient by the way of Equation (3) is also a function of the GMPE used. Furthermore, Equation (3) provides the point-estimate of the correlation coefficient, which in reality contains uncertainty due to the finite number of recordings (i.e.  $N_{\text{record}}$ ) used in its determination.

Consideration and quantification of the uncertainty in  $\rho_{\ln IM_i, \ln IM_j}$  is complicated by the fact that it has non-constant variance. This means that, for example, if 20 data have a (point-estimated) correlation of 0.9 the 90% confidence of this estimate will be [0.79, 0.95], while if the (point-estimated) correlation is 0.4 the 90% confidence interval will be [0.025, 0.68]. That is, the closer the absolute value of the correlation coefficient is to 1.0, the smaller the standard deviation in the estimated correlation is. The most effective means to correctly account for this non-constant variance is to use a transformation of the correlation coefficient. The Fisher  $z$  transformation [28, 29] is an approximate variance stabilizing transformation of the correlation coefficient given by:

$$z = \frac{1}{2} \ln \left( \frac{1+\rho}{1-\rho} \right) = \tanh^{-1}(\rho) \quad (4)$$

where  $\rho$  is the Pearson correlation coefficient,  $\ln$  is the natural logarithm;  $\tanh^{-1}$  is the inverse hyperbolic tangent function; and  $z$  is the transformed correlation coefficient. It can be shown that  $z$  has approximately a normal distribution with mean given by Equation (4), and variance given by:

$$\text{Var}[z] = \frac{1}{N-3} \quad (5)$$

where  $N$  is the sample size. The convergence of  $z$  to a normal distribution is very rapid (as  $N$  increases) [28] for the practical purposes as used herein. An alternative to the use of Equation (5) for estimating the variance in  $z$ , which is utilized herein and also applicable for all sample sizes, is the bootstrap method [26].

The use of Equation (5), or the bootstrap method, provides the uncertainty in the point-estimate of the correlation coefficient due to the finite number of samples used in the computation of Equation (3). However, it was also mentioned that the estimated correlation coefficient is a function of the GMPEs used, and therefore that there is uncertainty in the selection of which GMPE should be used. This uncertainty is accounted by using the four aforementioned GMPEs developed for active shallow crustal tectonic regions. It is assumed that each GMPE is equally valid, meaning that each is given a logic tree weight of 0.25 [30]. It is clearly acknowledged that there are several problems associated with the logic tree formalism for handling model uncertainty [31–33], however in the absence of a simple practical alternative, the logic tree formalism is adopted nonetheless.

Given that the transformed correlation coefficient is adequately represented by a normal distribution (the validity of this in light of GMPE uncertainty will be discussed subsequently), it follows that the median value of the correlation coefficient,  $\rho_{50}$ , is equal to the inverse Fisher transformation of the mean of  $z$ ,  $\mu_z$ , specifically:

$$\rho_{50} = \left( \frac{e^{2\mu_z} - 1}{e^{2\mu_z} + 1} \right) = \tanh(\mu_z) \quad (6)$$

While it may initially seem complicated to have uncertainty in a correlation coefficient (which defines the linear dependence between the uncertainties in two variables), it is worth noting that it is the same as the idea of having logic trees to represent uncertainty in the standard deviation of GMPEs. Over time such correlation equation uncertainty is likely to become an indispensable part of any model, in the same way that PSHA was first proposed without the consideration of

uncertainties in ground motion prediction [34], while nowadays a seismic hazard analysis is not considered to be a PSHA if it does not consider ground motion prediction uncertainty (e.g. [35]).

## EMPIRICAL CORRELATION RESULTS

### *Median correlation coefficient*

Figure 1 illustrates the normalized residuals for two different intensity measure combinations obtained from (a single bootstrap sample of) the adopted data set using the BA08 GMPE. Figure 1(a) illustrates there is a strong correlation between the residuals of PGA and ASI, while there is only a moderate correlation between the residuals of PGA and SI.

Figure 2 illustrates the correlation coefficients computed for three of the four different intensity measures examined (i.e. for ASI, SI, and PGA). In each of Figure 2(a)–(c) box plots [26] are used to illustrate the distribution of the correlation coefficient due to sample size uncertainty (obtained from bootstrap sampling), whereas the multiple box plots illustrate the effect of which GMPE is used.

As one would expect, given that ASI represents an average measure of high-frequency ground motion intensity, then it has a high correlation with PGA (Figure 2(b)). On the other hand, as SI represents an average measure of moderate period ground motion intensity, then it has a lower correlation with ASI and PGA (both high-frequency IMs).

It can be seen that variability in the correlation coefficients (as indicated by the ‘whiskers’ of the box plots) due to sample uncertainty is approximately 0.07, 0.025, and 0.08 for  $\rho_{\ln \text{ASI}, \ln \text{SI}}$ ,  $\rho_{\ln \text{ASI}, \ln \text{PGA}}$ , and  $\rho_{\ln \text{SI}, \ln \text{PGA}}$ , respectively. Thus, it can be clearly seen that the uncertainty in the

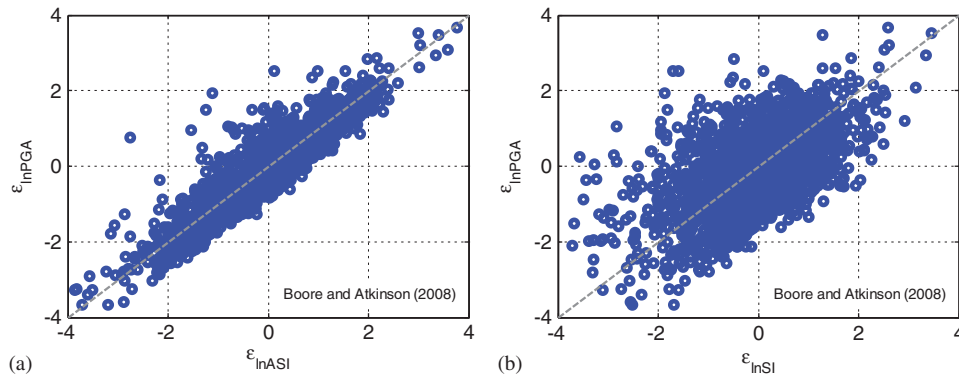


Figure 1. Examples of the correlations obtained between the residuals of: (a) PGA and ASI; and (b) PGA and SI.

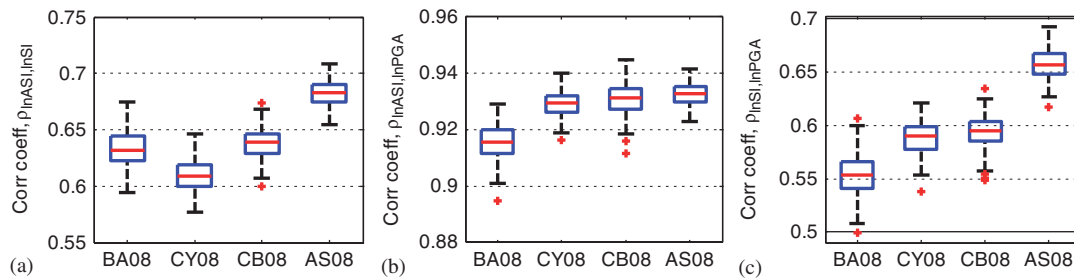


Figure 2. Correlation coefficients obtained between three different ground motion intensity measure pairs: (a) ASI and SI; (b) ASI and PGA; and (c) SI and PGA. In each figure box plots represent the uncertainty due to the finite sample size, whereas the different box plots illustrate the uncertainty due to the selected GMPE.

correlation coefficient due to sample size uncertainty is a function of the correlation coefficient value, and hence the benefit of the aforementioned Fisher  $z$  transformation which will be utilized subsequently (i.e. Equation (4)). It can also be seen that there is a difference between the correlation coefficients estimated using the four different GMPEs, which broadly speaking is of the same order of magnitude as the uncertainty in the correlation coefficient due to sample size uncertainty.

Figures 3(a)–(c) illustrate the empirical correlations obtained between SA at various vibration periods and ASI, SI, and PGA, respectively. In these figures, the median and 90% confidence interval of the correlation coefficient for a given GMPE (i.e. considering only sample size uncertainty) are indicated by solid and dashed lines, respectively. Different line colours are used to represent the distribution of the correlation coefficient obtained using the four different GMPEs considered. Based on similarities in the frequency range which the different IMs represent it can be intuitively appreciated that ASI is highly correlated with spectral accelerations in the range of 0.1–0.5 s (over which ASI is defined), and PGA is highly correlated with high-frequency spectral ordinates. In order to understand the empirical correlation of SI and SA it is useful to recall that SI is defined as the integral of the pseudo spectral velocity from 0.1–2.5 s, and that spectral velocity ordinates generally increase appreciably with increasing period over this range. Hence, the majority of the magnitude of SI for a single ground motion comes from the spectral velocities at these longer periods (i.e. 0.7–2.5 s). Hence, it can be seen in Figure 3(b) that SI is strongly correlated with spectral acceleration ordinates over this moderate period range.

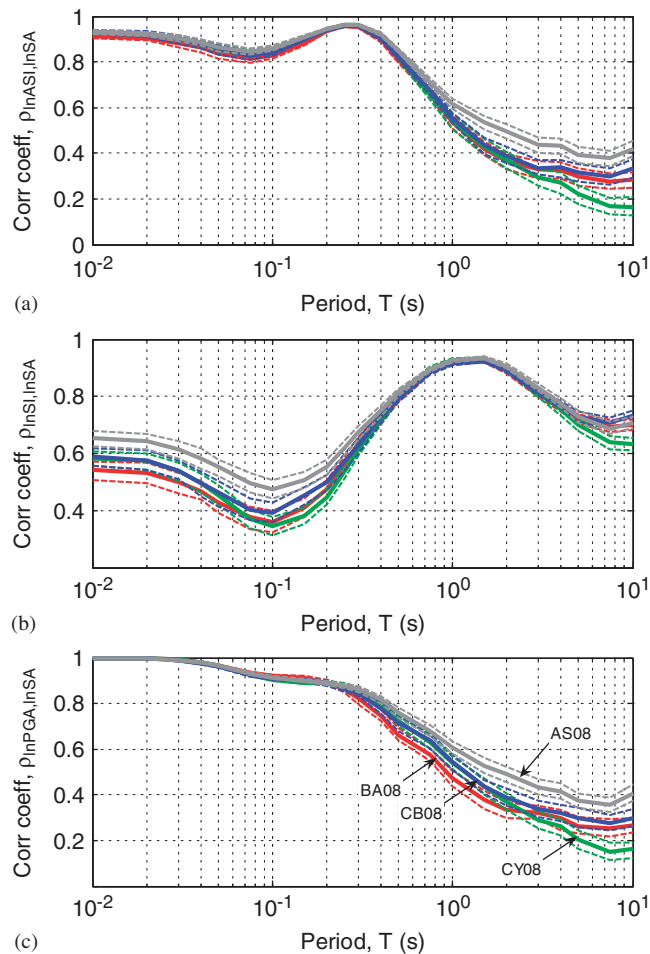


Figure 3. Correlation coefficients obtained between SA and (a) ASI; (b) SI; and (c) PGA. In each figure the solid and dashed lines represent the median and 90% confidence interval due to finite sample size, while the different coloured lines illustrate the uncertainty due to the selected GMPE.

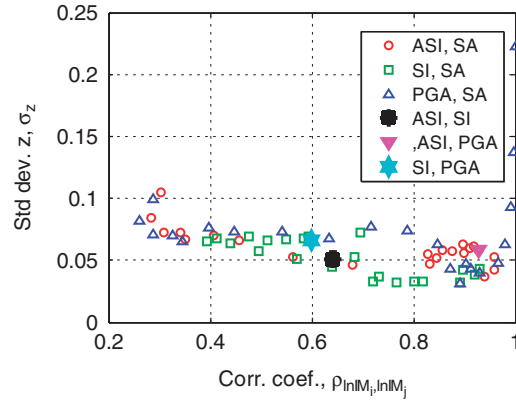


Figure 4. The standard deviation in the transformed correlation coefficient for the various intensity measure pairs considered.

For both  $\rho_{\ln \text{ASI}, \ln \text{SA}}$  and  $\rho_{\ln \text{SI}, \ln \text{SA}}$ , it can be seen that there is an increase in correlation coefficient for very high-frequency ground motion (i.e. going from  $T=0.1$  to  $T=0.01$ , both  $\rho_{\ln \text{ASI}, \ln \text{SA}}$ , and  $\rho_{\ln \text{SI}, \ln \text{SA}}$  increase), something which has also been observed by Baker and Jayaram [8, Figure 2]. Since, as the vibration period tends to zero, the value of SA tends to PGA, then it follows that this observed increase is the result of  $\rho_{\ln \text{ASI}, \ln \text{SA}}$ , and  $\rho_{\ln \text{SI}, \ln \text{SA}}$  tending toward  $\rho_{\ln \text{ASI}, \ln \text{PGA}}$ , and  $\rho_{\ln \text{SI}, \ln \text{PGA}}$ , respectively.

#### *Standard deviation of transformed correlation coefficient*

The previous section examined primarily the magnitude of the correlation coefficient between various IMs. As it is intended to develop correlation equations for both the median correlation coefficient and its uncertainty, then it is useful to examine quantitatively some aspects of the uncertainty in the transformed correlation coefficient. Figure 4 illustrates the empirical relationship between the median value of the correlation coefficient, and the standard deviation in the transformed correlation coefficient (i.e. including both sample size and GMPE uncertainty). First, it can be seen that apart from a few exceptions,  $\sigma_z$  is in the range [0.03, 0.09]. Second, given that a total of 1842 ground motions were used in this study, Equation (5) can be used to ascertain that the approximate portion of  $\sigma_z$  due to sample size uncertainty is  $1/\sqrt{1842-3}=0.023$ . Correlations involving spectral accelerations at longer periods used a slightly smaller number of ground motions (because of the maximum useable period for each ground motion), for example, 802 motions were used for correlations involving  $\text{SA}(T=7.5\text{s})$  giving a sample size uncertainty of 0.035. Thus, the magnitude of the sample size uncertainty indicates that the total standard deviation (i.e. comprising sample size and GMPE uncertainty),  $\sigma_z$ , in the transformed correlation coefficient, is composed largely by GMPE uncertainty and to a lesser extent sample size uncertainty.

## PARAMETRIC APPROXIMATIONS

#### *Adequacy of the normal distribution approximation*

It was previously mentioned that the uncertainty in  $z$  due to sample size was normally distributed. When considering also uncertainty in  $z$  due to GMPE uncertainty the assumption that  $z$  has a normal distribution may no longer be strictly appropriate. Figure 5 illustrates the empirical distribution of the transformed correlation,  $z$ , between ASI and SI (i.e. Figure 2(a)). Also shown is the normal distribution based on the point estimated mean and standard deviation of  $z$ , as well as the Lilliefors [26, 36] goodness of fit bounds for the 10% significance level. The fact that the empirical distribution intersects these goodness of fit bounds signifies that the assumption of normality for  $z$  can be rejected at the 10% significance level.

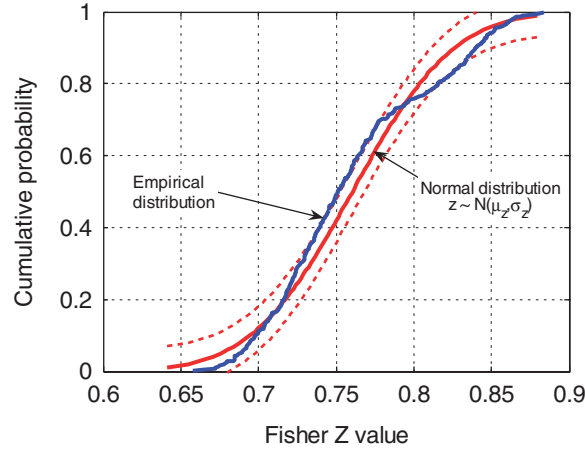


Figure 5. Illustration of the adequacy of the approximation that  $z$  has a normal distribution due to sample size and GMPE uncertainty for the correlation between SI and ASI.

One important thing to bear in mind in the examination of Figure 5 is that the empirical distribution of  $z$  has been obtained based on the logic tree (discrete) treatment of GMPE uncertainty. It is likely that the four different GMPEs used here do not provide a collective exhaustive representation of GMPE uncertainty [31, 32], and it is speculated that a more collectively exhaustive representation is likely to lead to a ‘smoother’ empirical distribution, for which the normal distribution may be more appropriate. Irrespective of whether the above speculation is correct or not it is timely to note that as the significance of ground motion correlation uncertainty, in ground motion selection, for example, has not yet been examined the aim here is, first and foremost, to adequately represent the magnitude of the standard deviation in the (transformed) correlation coefficient with less attention to the details of its specific distribution (i.e. in line with the principle of consistent crudeness [37]). Therefore, it is assumed herein that  $z$  can be adequately represented by the normal distribution, which has the benefit of being uniquely defined by its mean,  $\mu_z$ , and standard deviation,  $\sigma_z$ .

#### *Format for presentation of correlation equations*

As uncertainty in ground motion correlation equations has not been considered in applications which utilize such equations (e.g. ground motion selection) then it is desired to develop the parametric correlation equations such that they are simple to apply if one does not want to explicitly consider such correlation uncertainty. Therefore, rather than developing equations for the mean and standard deviation of  $z$  directly, use is made of Equation (1). Specifically, parametric equations are developed for: (i) the median correlation coefficient,  $\rho_{50}$ , and (ii) the standard deviation of the transformed correlation coefficient,  $\sigma_z$ . In this manner, if an analyst wishes to neglect correlation coefficient uncertainty they can simply use  $\rho_{50}$  directly. On the other hand, if one wants to explicitly account for correlation coefficient uncertainty then Equation (6) can be used to determine  $\mu_z$  from the equation for  $\rho_{50}$ , and then with  $\sigma_z$ , the correlation coefficient for a given  $x$ th percentile can be obtained from:

$$\rho_x = \left( \frac{e^{2z_x} - 1}{e^{2z_x} + 1} \right) = \tanh(z_x) \quad (7)$$

where  $z_x$  is the  $x$ th percentile of a normal distribution with mean and standard deviation  $\mu_z$  and  $\sigma_z$ , respectively [26].

#### *Correlation equations between ASI, SI, and PGA*

Table I provides the median correlation coefficient,  $\rho_{50}$ , and the standard deviation of the transformed correlation,  $\sigma_z$ , between ASI, SI, and PGA. These values have been obtained directly from



Table I. Proposed median correlation,  $\rho_{50}$ , and the standard deviation of the transformed correlation,  $\sigma_z$  (shown in brackets) between ASI, SI, and PGA.

$\rho_{50}$ ( $\sigma_z$ )	ASI	SI	PGA
ASI	1.0	0.641 (0.051)	0.928 (0.058)
SI	—	1.0	0.599 (0.066)
PGA	Symm.	—	1.0

the mean and standard deviation of the  $z$  values obtained from the aforementioned correlation analyses (i.e. Figure 2). No smoothing of the  $\rho_{50}$  or  $\sigma_z$  values have been conducted.

#### *Correlation equations for SA with ASI, SI, and PGA*

The correlation coefficients computed involving SA have 23 different values for the 23 vibration periods between  $T = [0.01, 10]$  for which SA was computed. Rather than providing the median and standard deviation values of the correlation coefficient for each of these 23 different vibration periods it is beneficial to represent the correlation coefficient,  $\rho_{\ln \text{IM}_i, \ln \text{SA}(T)}$  (where  $\text{IM}_i$  is one of ASI, SI, or PGA), as a continuous function of vibration period,  $T$ . Such a continuous function offers the benefits of: (i) interpolation of the correlation coefficient at vibration periods different than the 23 periods for which it was explicitly computed; and (ii) a smooth variation of correlation coefficient as a function of vibration period, rather than the more ‘jagged’ values that result due to the finite number of ground motion data, and specific GMPEs used.

The specific functional form used to represent the variation in the (median) correlation coefficient with vibration period is a piecewise function with each piecewise segment having the following functional form:

$$\rho_{\ln \text{IM}_i, \ln \text{SA}(T)} = \frac{a_n + b_n}{2} - \frac{a_n - b_n}{2} \tanh[d_n \ln(T/c_n)], \quad e_{n-1} \leq T < e_n \quad (8)$$

where  $\tanh$  is the hyperbolic tangent function;  $a_n$ ,  $b_n$ ,  $c_n$ , and  $d_n$  are empirical constants used to fit the observed empirical correlation coefficient for piecewise segment  $n$ ; and  $e_n$  defines the period range for each of the piecewise segments. Generally the vibration period range  $T = [0.01, 10]$  is covered by either two or three separate piecewise segments of the form of Equation (8). The specific functional form given by Equation (8) has been selected because it is both flexible and also its parameters can be understood physically, making the fitting of these parameters very simple. First, it is noted that the hyperbolic tangent function varies from  $[-1, 1]$ . Therefore, the parameter  $a_n$  effectively represents the correlation between  $\text{IM}_i$  and  $\text{SA}(T = e_{n-1})$ , whereas the parameter  $b_n$  effectively represents the correlation between  $\text{IM}_i$  and  $\text{SA}(T = e_n)$ . The parameter  $c_n$  defines the vibration period at which the correlation coefficient between  $\text{IM}_i$  and  $\text{SA}(T = c_n)$  is equal to the average correlation value over the range  $e_{n-1} < T < e_n$ . (i.e.  $\rho_{\ln \text{IM}_i, \ln \text{SA}(T=c_n)} = (a_n + b_n)/2$ ); finally the parameter  $d_n$  controls the slope of the hyperbolic tangent function (e.g. how quickly the correlation approaches  $a_n$  as  $T$  decreases from  $T = c_n$  toward  $T = e_{n-1}$ ).

In order to fit the hyperbolic tangent function given by Equation (8) in a piecewise manner to the empirically observed correlation data, it was necessary to specify the ends of the piecewise segments (i.e. the  $e_n$ 's). As Equation (8) is either monotonically increasing or decreasing (depending on the values  $a_n$  and  $b_n$ ) then these piecewise segments were simply chosen as points of maxima or minima in the empirical data. Equation (8) was then fit using least squares to minimize the misfit in the mean transformed  $z$  value (i.e. to minimize the error between the Fisher transformation of Equation (8) and the empirical values of  $\mu_z$ ). It is important to conduct least squares on the transformed  $z$  values as they have approximately constant variance (which the correlation coefficient values do not), making the least squares solution more optimal. The parameters of each piecewise segment were also constrained so that the parametric equation is a continuous piecewise function of  $T$ .

The obtained parametric equations were then plotted and compared with the empirical correlation values, and adjusted manually if necessary.

Given the unknown significance of the magnitude of the standard deviation in the transformed correlation,  $\sigma_z$  (because such a correlation coefficient uncertainty has, as far as the author is aware, not been considered in any applications), it is desired to represent it in a simplified, yet sufficient manner. As a result, piecewise linear relationships between  $\sigma_z$  and  $\ln(T)$  were used, specifically:

$$\sigma_z(\ln \text{ ASI}, \ln \text{ SA}) = \begin{cases} 0.05, & 0.01 \leq T < 0.8 \\ 0.05 + 0.015 \ln(T/0.8), & 0.8 \leq T \leq 10 \end{cases} \quad (9)$$

$$\sigma_z(\ln \text{ SI}, \ln \text{ SA}) = \begin{cases} 0.07 - 0.0018 \ln(T/0.01), & 0.01 \leq T < 0.15 \\ 0.065 - 0.0285 \ln(T/0.15), & 0.15 \leq T < 0.4 \\ 0.037, & 0.4 \leq T < 6 \\ 0.037 + 0.046 \ln(T/6), & 6 \leq T \leq 10 \end{cases} \quad (10)$$

$$\sigma_z(\ln \text{ PGA}, \ln \text{ SA}) = \begin{cases} 0.22 - 0.112 \ln(T/0.01), & 0.01 \leq T < 0.05 \\ 0.04 + 0.0085 \ln(T/0.05), & 0.05 \leq T \leq 10 \end{cases} \quad (11)$$

Tables II–IV provide the numerical values of the parameters in Equation (8) for the median correlations of ASI, SI, and PGA with  $\text{SA}(T)$ , respectively. For  $\rho_{\ln \text{ ASI}, \ln \text{ SA}(T)}$  and  $\rho_{\ln \text{ SI}, \ln \text{ SA}(T)}$  three piecewise segments were used, while for  $\rho_{\ln \text{ PGA}, \ln \text{ SA}(T)}$  only two segments were necessary. Figure 6 illustrates a comparison between the parametric correlation equations (i.e. Equation (8) and Equations (9)–(11)) and the empirical correlation values obtained from the aforementioned analyses. In each of the figures, the solid line represents the median correlation coefficient,  $\rho_{50}$ , whereas the dashed lines represent the 16th and 84th percentiles. It can be seen that, in general, the

Table II. Parameters in Equation (8) defining the piecewise variation of the median correlation between ASI and  $\text{SA}(T)$ ,  $\rho_{\ln \text{ ASI}, \ln \text{ SA}(T)}$ .

$n$	$e_n$	$a_n$	$b_n$	$c_n$	$d_n$
0	0.01	—	—	—	—
1	0.075	0.927	0.823	0.04	1.8
2	0.3	0.823	0.962	0.14	2.2
3	10	1.05	0.29	0.80	1.0

Table III. Parameters in Equation (8) defining the piecewise variation of the median correlation between SI and  $\text{SA}(T)$ ,  $\rho_{\ln \text{ SI}, \ln \text{ SA}(T)}$ .

$n$	$e_n$	$a_n$	$b_n$	$c_n$	$d_n$
0	0.01	—	—	—	—
1	0.1	0.60	0.38	0.045	1.5
2	1.4	0.38	0.94	0.33	1.4
3	10	0.95	0.68	3.1	1.6

Table IV. Parameters in Equation (8) defining the piecewise variation of the median correlation between PGA and  $\text{SA}(T)$ ,  $\rho_{\ln \text{ PGA}, \ln \text{ SA}(T)}$ .

$n$	$e_n$	$a_n$	$b_n$	$c_n$	$d_n$
0	0.01	—	—	—	—
1	0.2	1.00	0.895	0.06	1.6
2	10	0.97	0.25	0.80	0.8

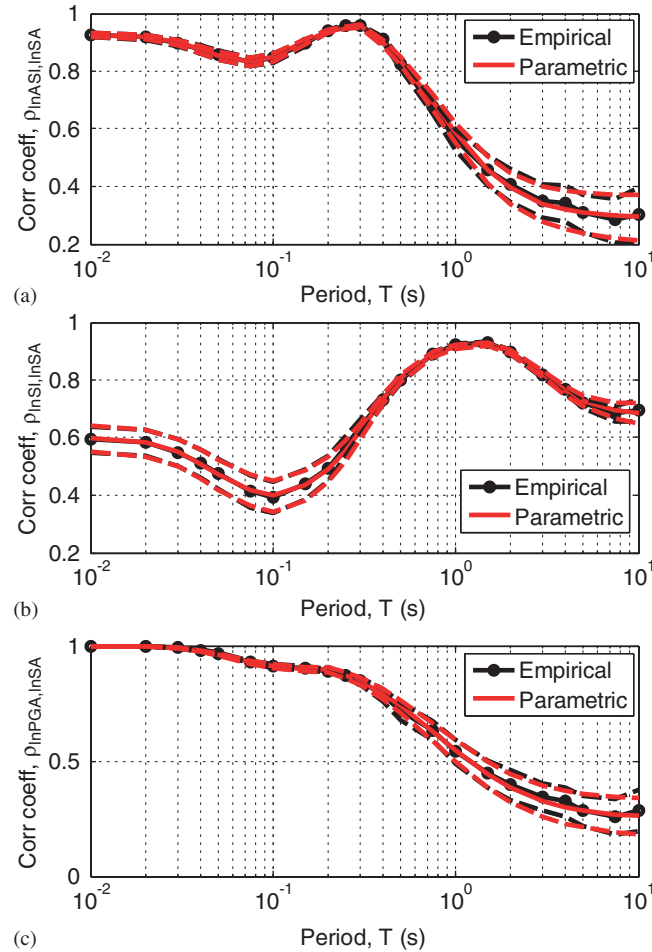


Figure 6. Comparisons of the empirical correlations and parametric fit of (a) ASI; (b) SI; and (c) PGA with spectral accelerations (SAs). Solid lines show the median, while dashed lines represent the 16th and 84th percentiles.

median correlation coefficient is very well approximated by Equation (8). Based on the examination of the 16th and 84th percentiles it also appears that Equations (9)–(11) are sufficient to characterize the magnitude of  $\sigma_z$ .

Figure 7 compares explicitly the empirical values of  $\sigma_z$  with the parametric fit provided by Equations (9)–(11). It can be seen that generally the uncertainty in the transformed correlation coefficient increases at very long periods as a result of a significant number of the ground motion records having a maximum usable period lower than these vibration periods. Hence, there are both less samples used when computing these correlation coefficient (leading to larger sample size uncertainty), and also the lower number of ground motions used provide less constraint for GMPEs leading also to larger GMPE uncertainty.

#### Comparison with other correlations between PGA and SA

While empirical correlation equations for five of the six IM combinations have not, as far as the author is aware, been previously developed, there are several previous instances where correlations between PGA and spectral ordinates have been computed. The first of these is the empirical  $\rho_{\ln \text{PGA}, \ln \text{SA}(T)}$  prediction equation developed by Baker [10]. Baker [10] used a subset of 517 ground motions from an older version of the NGA database used in this study, and the GMPEs of Abrahamson and Silva [38] and Boore *et al.* [39] for computing the distributions of SA and PGA,

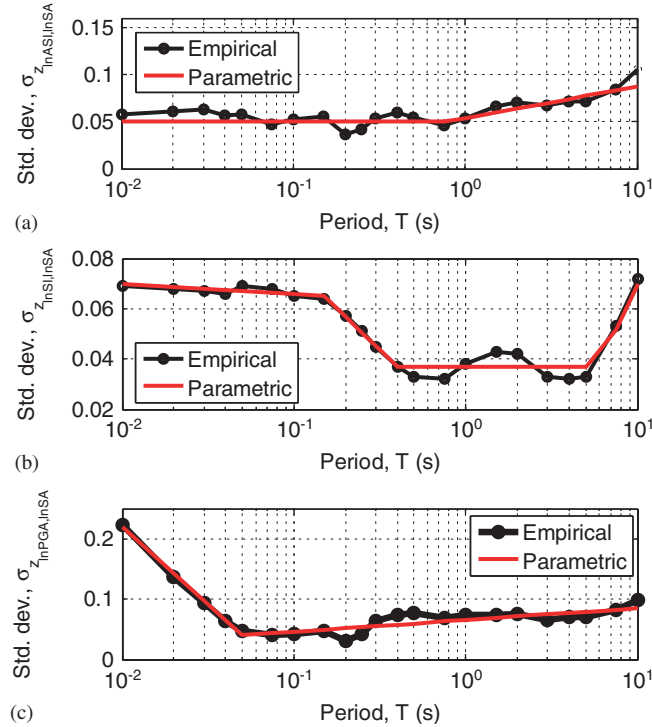


Figure 7. Comparisons of the standard deviation of the transformed correlation coefficient from the ground motions analysed and the proposed parametric fit for: (a) ASI, SA; (b) SI, SA; and (c) PGA, SA.

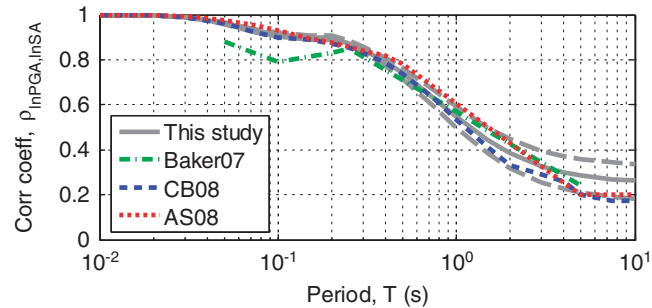


Figure 8. Comparison of the parametric correlation model for SA and PGA developed in this study with previous models.

respectively. This equation was valid over the period range  $T = [0.05, 5]$ . The second and third instances of correlations between PGA and SA being presented are in the intra-event correlation values tabulated as part of the CB08 and AS08 GMPEs. Both CB08 and AS08 use the PGA on bedrock in order to determine the non-linear scaling of site response in their SA GMPEs. As such, when computing the standard deviation of SA using these equations one needs the correlation between PGA (on rock) and SA at the vibration period of interest. The CB08 and AS08 models provide tabulated values for spectral periods over the range  $T = [0.01, 10]$ .

Figure 8 compares the median, 16th and 84th percentiles of the  $\rho_{\ln PGA, \ln SA(T)}$  model developed here with the model developed by Baker [10] and the tabulated correlation values in CB08 and AS08. It can be seen that, in general, all of the models are in agreement. Two specific points are worthy of note with regard to Figure 8. First, for vibration periods less than  $T = 0.3$  s the Baker07 model correlation values initially decrease, in comparison with the model developed in this study and the CB08 and AS08 values, which increase monotonically with decreasing period

to correlations of 1.0 at very short periods. Second, over all vibration periods the CB08 and AS08 models (and the Baker07 model for  $T > 0.03s$ ) generally lie 'inside' the 16th and 84th percentile values of the model developed in this study, indicating their consistency with the present model.

## DISCUSSION

The empirical correlation equations developed in this paper are based on a database of 1842 ground motions from the NGA database, and the four aforementioned GMPEs (i.e. BA08, CY08, CB08, AS08) developed using ground motions from the NGA database. Various studies have demonstrated that the NGA GMPEs (with possibly minor modifications) cannot be statistically rejected as appropriate for predicting PGA and SA amplitudes for various active shallow crustal tectonic regions throughout the world [40–45]. This result also suggests that the underlying NGA database, and consequently the empirical correlation equations developed in this paper, can also be considered generically applicable for such active shallow crustal tectonic regions.

## CONCLUSIONS

This paper has examined, and developed parametric models for, the empirical correlation between PGA, spectral acceleration, spectrum intensity, and acceleration spectrum intensity. The correlation equations are considered applicable for shallow crustal earthquakes and were developed using the Next Generation Attenuation (NGA) ground motion database, and four of the NGA GMPEs. The magnitude of the observed correlations between the various intensity measures could be explained simply based on the frequency ranges of ground motion which the intensity measures represent. Explicit consideration was given to the epistemic uncertainty in the correlation equations due to both the adopted ground motion database and ground motion prediction equations, and as such, correlation equations were developed to provide both the median and standard deviation of the correlation coefficient. The resulting correlation equations enable the joint consideration of these four ground motion intensity measures in various performance-based earthquake engineering applications, including in ground motion selection using frameworks such as the generalized conditional intensity measure (GCIM) approach.

## REFERENCES

1. Bradley BA, Cubrinovski M, Dhakal RP, MacRae GA. Intensity measures for the seismic response of pile foundations. *Soil Dynamics and Earthquake Engineering* 2009; **29**(6):1046–1058.
2. Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA. Prediction of spatially distributed seismic demands in structures: ground motion and structural response. *Earthquake Engineering and Structural Dynamics* 2009; **39**(5):501–520.
3. Shome N, Cornell CA, Bazzurro P, Carballo JE. Earthquakes, records, and nonlinear responses. *Earthquake Spectra* 1998; **14**(3):469–500.
4. Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. *Earthquake Spectra* 2007; **23**(2):357–392.
5. Bradley BA. A generalized conditional intensity measure approach and holistic ground motion selection. *Earthquake Engineering and Structural Dynamics* 2010; **39**(12):1321–1342.
6. Inoue T, Cornell CA. Seismic hazard analysis of multi-degree-of-freedom structures, 1990; 70.
7. Baker JW, Cornell CA. Correlation of response spectral values for multi-component ground motions. *Bulletin of the Seismological Society of America* 2006; **96**(1):215–227.
8. Baker JW, Jayaram N. Correlation of spectral acceleration values from NGA ground motion models. *Earthquake Spectra* 2008; **24**(1):299–317.
9. Arias A. A measure of earthquake intensity. In *Seismic Design for Nuclear Power Plants*, Hansen RJ (ed.). MIT Press: Cambridge, MA, 1970; 438–483.
10. Baker JW. Correlation of ground motion intensity parameters used for predicting structural and geotechnical response. *Tenth International Conference on Application of Statistics and Probability in Civil Engineering*, Tokyo, Japan, 2007; 8.
11. Housner GW. Spectrum intensities of strong-motion earthquakes. *Symposium on Earthquakes and Blast Effects on Structures*. Earthquake Engineering Research Institute: Los Angeles, CA, 1952.

12. Von Thun J, Roehm L, Scott G, Wilson J. Earthquake ground motions for design and analysis of dams. *Earthquake Engineering and Soil Dynamics II—Recent Advances in Ground-motion Evaluation, Geotechnical Special Publication* 1988; **20**:463–481.
13. Chiou B, Darragh R, Gregor N, Silva WJ. NGA project strong-motion database. *Earthquake Spectra* 2008; **24**(1):23–44.
14. NZS 1170.5. 2004. Structural design actions, Part 5: Earthquake actions, New Zealand, Standards New Zealand.
15. Priestley MJN, Calvi GM, Kowalsky MJ. *Displacement-based Seismic Design of Structures*. IUSS Press: Pavia, Italy, 2007; 721.
16. Haselton CB, Baker JW, Bozorgnia Y, Goulet CA, Kalkan E, Luco N, Shantz T, Shome N, Stewart JP, Tothong P, Watson-Lamprey J, Zareian F. Evaluation of ground motion selection and modification methods: predicting median interstorey drift response of buildings. *PEER Technical Report 2009/01*, 2009; 288.
17. ASCE. 2006. Minimum design loads for buildings and other structures. American Society of Civil Engineers, *ASCE Standard No. 007-05*.
18. Bradley BA, Dhakal RP, Cubrinovski M, MacRae GA. Prediction of spatially distributed seismic demands in structures: from structural response to loss estimation. *Earthquake Engineering and Structural Dynamics* 2009; **39**(6):591–613.
19. Bradley BA. Site specific and spatially distributed prediction of acceleration spectrum intensity. *Bulletin of the Seismological Society of America* 2009; **100**(2):792–801.
20. Bradley BA, Cubrinovski M, MacRae GA, Dhakal RP. Ground motion prediction equation for spectrum intensity from spectral acceleration relationships. *Bulletin of the Seismological Society of America* 2009; **99**(1):277–285.
21. Boore DM, Atkinson GM. Ground-motion prediction equations for the average horizontal component of PGA, PGV, and 5%-damped PSA at spectral periods between 0.01s and 10.0s. *Earthquake Spectra* 2008; **24**(1):99–138.
22. Chiou BSI, Youngs RR. An NGA Model for the average horizontal component of peak ground motion and response spectra. *Earthquake Spectra* 2008; **24**(1):173–215.
23. Campbell KW, Bozorgnia Y. NGA ground motion model for the geometric mean horizontal component of PGA, PGV, and 5% damped linear elastic response spectra for periods ranging from 0.01 to 10s. *Earthquake Spectra* 2008; **24**(1):139–171.
24. Abrahamson NA, Silva WJ. Summary of the Abrahamson & Silva NGA ground motion relations. *Earthquake Spectra* 2008; **24**(1):67–97.
25. Chiou B, Youngs RR. Chiou and Youngs PEER-NGA empirical ground motion model for the average horizontal component of peak acceleration and Pseudo-Spectral acceleration for spectral periods of 0.01 to 10 Seconds. 2006; 219.
26. Ang AHS, Tang WH. *Probability Concepts in Engineering: Emphasis on Applications in Civil and Environmental Engineering*. Wiley: New York, 2007.
27. Abrahamson NA, Youngs RR. A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America* 1992; **82**(1):505–510.
28. Kutner M, Nachtsheim C, Neter J, Li W. *Applied Linear Statistical Models*. McGraw-Hill/Irwin: New York, 2005; 1396.
29. Fisher R. Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population. *Biometrika* 1915; **10**(4):507–521.
30. Kulkarni RB, Youngs RR, Coppersmith KJ. Assessment of confidence intervals for results of seismic hazard analysis. *Eighth World Conference on Earthquake Engineering*, San Francisco, CA, 1984; 263–270.
31. Bommer JJ, Scherbaum F. The use and misuse of logic trees in probabilistic seismic hazard analysis. *Earthquake Spectra* 2008; **24**(4):997–1009.
32. Abrahamson NA, Bommer JJ. Probability and uncertainty in seismic hazard analysis. *Earthquake Spectra* 2005; **21**(2):603–607.
33. Page MT, Carlson JM. Methodologies for earthquake hazard assessment: model uncertainty and the WGCEP-2002 forecast. *Bulletin of the Seismological Society of America* 2006; **96**(5):1624–1633.
34. Cornell CA. Engineering seismic risk analysis. *Bulletin of the Seismological Society of America* 1968; **58**(5):1583–1606.
35. Bommer JJ, Abrahamson NA. Why do modern probabilistic seismic-hazard analyses often lead to increased hazard estimates? *Bulletin of the Seismological Society of America* 2006; **96**(6):1967–1977.
36. Lilliefors H. On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association* 1967; **62**:399–402.
37. Elms DG. Consistent crudeness in system construction. In *Optimisation and Artificial Intelligence in Civil Engineering*, Topping BHV (ed.). Kluwer Academic Publishers: Dordrecht, 1992; 71–85.
38. Abrahamson NA, Silva WJ. Empirical response spectral attenuation relations for shallow crustal earthquakes. *Seismological Research Letters* 1997; **68**(1):94–126.
39. Boore DM, Joyner WB, Fumal TE. Equations for estimating horizontal response spectra and peak acceleration from western North American earthquakes: a summary of recent work. *Seismological Research Letters* 1997; **68**(1):128–153.
40. Scasserra G, Stewart JP, Bazzurro P, Lanzo G, Mollaioli F. A comparison of NGA ground-motion prediction equations to Italian data. *Bulletin of the Seismological Society of America* 2009; **99**(5):2961–2978.
41. Stafford PJ, Strasser FO, Bommer JJ. An evaluation of the applicability of the NGA models to ground-motion prediction in the Euro-Mediterranean region. *Bulletin of Earthquake Engineering* 2008; **6**:149–177.

42. Atkinson GM. Ground-motion prediction equations for Hawaii from a referenced empirical approach. *Bulletin of the Seismological Society of America* 2010; **100**(2):751–761.
43. Campbell KW, Bozorgnia Y. Next generation attenuation (NGA) empirical ground motion models: can they be used in Europe? *First European Conference on Earthquake Engineering and Seismology*, Geneva, Switzerland, 2006; 10.
44. Shoja-Taheri J, Naserieh S, Hadi G. A test of the applicability of NGA models to the strong ground-motion data in the Iranian plateau. *Journal of Earthquake Engineering* 2010; **14**(2):278–292.
45. Peruš I, Fajfar P. Ground-motion prediction by a non-parametric approach. *Earthquake Engineering and Structural Dynamics* 2010; **39**(12):1395–1416.